A Brief Survey on Deep Learning-Based Brain Tumour Detection Techniques

Komma Thrivikram[1] & Dr.T.Ananth Kumar[2]

[1,2]*Department of Computer Science and Engineering, IFET College of Engineering, Villupuram, Tamilnadu, India.*

ABSTRACT

In recent years, the application of deep learning has demonstrated significant progress in various scientific subfields. Compared to other cutting-edge methods of processing and analysing images, deep learning algorithms performed significantly better. When applied in areas such as self-driving cars, where deep learning has been utilized, and the results are the best and most up-to-date currently available. In some situations, such as recognizing objects and playing games, deep learning performed significantly better than people did. One more industry that appears to have a lot to gain from deep learning is the medical field. There are a lot of patient records and data, and providing individualized care to each patient is becoming an increasingly important priority as a result. This indicates an immediate need for methods that are both efficient and reliable for processing and analyzing health informatics.
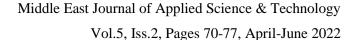
Keywords: Deep learning, MRI, DNN, Brain Tumour.

1. Introduction

Deep learning has been making significant contributions to various scientific fields in recent years, and these contributions have been widely recognised. When it came to several different tasks, such as the processing and analyzing images, deep learning methods proved to be superior to traditional ones [1]. In addition, deep learning excels at tasks that previously could not be completed automatically, such as driving autonomous vehicles. Deep learning performs significantly better than humans when it comes to specific tasks, such as recognizing objects and playing games. In addition to hospitals and private practices, information about patients can also be gathered using websites and apps designed specifically for mobile devices that are dedicated to healthcare. The process of automatically removing brain tumours from MRI scans is considered one of the most challenging problems in computer vision. Because, Deep Neural Networks (DNN) [2] are so effective at automatically separating images of brain tumours, numerous proposals investigate the ways in which they could be utilized in the image segmentation process. The training process for deeper neural networks typically takes a lot more time and requires a lot more computing power. This issue can be traced back to the gradient diffusion problem.

This paper demonstrates an automatic method based on Deep Residual Learning Networks for segmenting brain tumours into their respective parts (ResNet) [3]. Our objective was to find a solution that would circumvent DNN's gradient issue. ResNets have a higher level of accuracy than DNNs and can be trained much more quickly. The segmentation of brain tissue from multimodal MRI is typically considered to be the most critical step in the process of analyzing neuroimaging data.

The development of deep neural networks has contributed significantly to the rapid advancement of brain lesion segmentation over the past few years (DNNs). However, there are only a select few methods that can be used at the same time to differentiate between normal tissue and brain lesions. Typically, annotated datasets only focus on a single task and only use imaging protocols and modalities explicitly tailored to that task. Because of this, it is challenging to create a DNN that is capable of performing both tasks at the same time. This paper shows a new
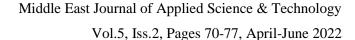
approach to building a joint tissue and lesion segmentation model by utilising task-specific hetero-modal domain-shifted and partially annotated datasets [4]. This approach aims to improve the accuracy of joint tissue and lesion segmentation. This technique is novel and distinctive in that it offers a fresh approach to the division of joint tissue and lesions. Beginning with a variation formulation of the joint problem, we show how the expected risk can be decomposed and empirically optimised in order to achieve the best possible outcome. This is accomplished by focusing on the relationship between the two variables. We make use of an upper risk bound whenever we are working with datasets that contain a variety of imaging modalities [5].

Over the last few decades, neurologists have devoted a significant amount of their time and energy to the study of brain tumours. Gliomas are a type of brain tumour that originates in cells that are referred to as glial cells. On a scale that ranges from one to four, grades three and four demonstrate cancerous brain tumours, whereas grades one and two demonstrate benign brain tumours. The term "magnetic resonance imaging," more commonly abbreviated as "MRI," is a well-known medical technique that can be applied to detect brain tumours. The fact that the device can effectively examine the soft tissues of the brain is a significant help to the investigatory efforts of medical professionals [6]. An MRI can produce images in the following weighting schemes: T1-weighted (T1w), T1-weighted with contrast enhancement (T1wc), T2-weighted (T2w), and fluid-attenuated inversion recovery (FLAIR). When looking at MRI images, it can be challenging to determine the location of the edges of a tumour because they frequently appear identical, and the intensity of the tumour can change even within the same sequence. This makes it difficult to diagnose. Convolutional neural networks, more commonly abbreviated as CNNs, were utilised in the process of multimodal MRI brain tumour segmentation.

The CNN model is a cutting-edge piece of technology that combines feature extraction and classification into a single cohesive model. This indicates that the forecast will be closer to the mark as a result. The authors of suggested using a deep CNN model with small 33 convolution kernels in order to segment MRI images of tumours. They made CNN's deeper by using smaller filter kernels and adding more convolutional layers in a cascade, but the responsiveness of the field was the same as when they used larger kernels. Additionally, CNN methods are notoriously difficult to train due to the fact that as the process continues, the gradient has a tendency to vanish. Research using magnetic resonance (MR) [7] has as its long-term objective the determination of a method for classifying distinct types of tissues, such as brain tissues and other types of tissues, according to how they are constructed. People had the notion of classifying tissues prior to the development of MR imaging by the T1 and T2 relaxation times of the respective tissues. Cancer is a medical condition in which cells in a specific body area grow out of control and in an uncontrolled manner. All of the abnormalities that are characteristic of tumours, including lumps, micro-taxonomies, and deformities, are due to the abnormal growth of cells within the tumour. An intracranial tumour is a subtype of the brain tumour that results when cells in the brain divide without being controlled. Intracranial tumours can be fatal. According to research conducted by the National Brain Tumour Society (NBTS), between 20 and 40 percent of cancers are associated with metastases in the brain. According to the graph that depicts diagnoses, the number of people who have been found to have the same condition ranges from 98,000 to 170,000. Imaging assists medical professionals in locating and diagnosing diseases and provides them with essential information regarding the human body that is necessary for clinical procedures.
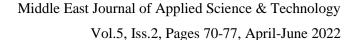
Additionally, having a strong understanding of the human body makes it simpler to communicate with surgical assistants regarding the appropriate actions to take during a procedure [8]. The acronym "MRI" stands for "magnetic resonance imaging," a type of medical imaging that creates detailed images of a patient's organs and tissues by employing a computer-generated magnetic field in conjunction with radio waves. This technique is abbreviated as "MRI." The majority of MRI machines have very powerful magnets that are shaped like tubes and are very long. The magnetic field of an MRI machine causes the water molecules in a person's body to move around for a brief period while the person is lying inside the machine. An MRI can produce images of the body's cross-sections because of the coordinated efforts of these atoms, which send out very small signals into the surrounding space [9].

2. Literature Survey

A method of machine learning referred to as Multi-Task Learning (also abbreviated as MTL), Multi-Task Learning has the objective of performing multiple tasks concurrently on a single dataset [10]. To achieve this goal, first it is necessary to devise a strategy for disseminating information or for demonstrating it. Next, a backend that is tailored to the activity must be developed. When a DNN is used for MTL, the initial layers of the network are typically shared among the different tasks, while the later layers of the network are trained specifically for each task. In the field of medical imaging, MTL has been utilised to combine segmentation with other tasks such as detection or classification with great success. The determination of the global loss function is accomplished by computing the weighted sum of the loss functions for each task. Kendall and Gal recently developed a method for estimating MTL loss weights that do not rely on Bayesian parameters. In 2018, the authors successfully implemented it in medical imaging using spatially adaptive task weighting and made it work. Even though the methods as mentioned above produce a variety of outputs from the same set of inputs, they do not model any kind of direct interaction between the outputs generated by the various tasks. It is possible to perform joint tissue and lesion segmentation in clinical settings; however, the assumption is typically made that the two outputs are conditionally independent of one another. Therefore, the issue of combining the results of all of these outputs into a single segmentation cannot be resolved using these methods. Also, MTL approaches such as do not have a way to deal with heteromodal datasets or differences in imaging properties that exist between task-specific databases.

Dealing with datasets collected in various environments requires a method known as domain adaptation, which is also known as "domain-shifted datasets" (DA). Learning a feature representation of the data that is not dependent on the domain is one of the more common approaches. Csurka (2017) recommended conducting an in-depth analysis of each of these different deep learning methods [11]. In response to a particular shift that has been observed, DA has formulated a variety of strategies and plans. For instance, data augmentation has been used to correct shifts in MR images that were brought about by a variety of MR bias fields (Sudre et al., 2017) or the presence of motion artifacts. In order to deal with missing modalities, Havaei et al. (2016) and Dorent et al. (2019a) have proposed network architectures that encode each modality into a shared modality-agnostic latent space. Recent findings suggest that it may be possible to learn a mapping between healthy and disease scans by employing CycleGANs or Variational Autoencoders. This possibility is supported by the findings of two separate research groups. These approaches were utilised in order to examine MRI scans of the brain. Although these strategies have
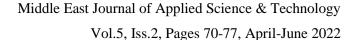
been successful in the past, how they function restricts their application to a particular kind of shift only. It has not been determined how to combine different acquisition protocols with different shift causes, such as determining whether or not there are lesions. This is a problem that has not been solved.

On the other hand, general DA approaches do not presuppose anything about the type of change they are attempting to investigate. These methodologies aim to bring the feature distributions of various domains into closer alignment with one another [12]. You can determine how dissimilar two distributions are to one another by using correlation distances or the maximum mean discrepancy. Both of these measures are available. On the other hand, more recent techniques are primarily driven by competition and have demonstrated significant potential in the field of medical imaging. On the other hand, the focus of these methods is almost entirely on the resolution of a single issue that applies to multiple domains.

Weakly-supervised, the process of learning, which is also known as WSL, is designed to function with annotations that are missing, incorrect, or not very precise. Because each task-specific dataset only contains one set of labels and the two sets cannot be combined, we can say that our problem is an example of learning with missing labels because each dataset only contains one set of labels. Li and Hoiem developed a method that allows a model that has been trained to perform one task to pick up and perform another task (2017) [13]. MTL and DA are combined in this approach through the utilisation of transfer learning. At the conclusion of the procedure, there are two models: one for the initial task, and another for the subsequent one. By employing a knowledge distillation loss, Kim et al. (2018) created a multitask model that is one of a kind. The objective here is to learn one task at a time while simultaneously keeping in mind the other task. Because of this, the WSL problem was recast as an MTL problem, which had the same restrictions for our use case as it did before. Based on the findings of this research, a novel approach to performing joint segmentation on datasets that are simultaneously task-specific, domain-shifted, and multimodal has been proposed.

Fully convolutional networks, also known as FCNs, are believed to be the first method to train a network completely from beginning to end. The process of semantic segmentation is carried out on a pixel-by-pixel basis using FCN. The FCN that is produced by first downsampling the network in layers and then employing an upsampling operator rather than a pooling operator [14]. The FCN architecture is created by combining the characteristics of the up-sampled layer with those of the corresponding down-sampled layer. Because the up-sampling component contains a large number of component channels, the network is able to transmit context data with a higher resolution. As a result, the path that is shown to contract (on the left) and the path that is shown to expand (on the right) are the same and make a U-shaped structure. The network will not function properly if the valid portion of each convolution is not utilised, and there are no layers that are completely connected.

The authors recommended making improvements to the U-net architecture model. They suggested substituting the convolutional blocks with residual blocks, similar to those described in the initial Unet. Each block consists of two convolutional units, each of which contains a Batch Normalization (BN) layer, an activation function, and a convolutional layer. Additionally, each unit contains a convolutional layer. PReLU, which is an abbreviation for "Parametric Rectified Linear Unit," is the term that is currently utilised to talk about the activation function [15].
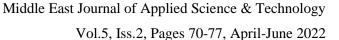
The ReLU function was incorporated into the model that was proposed. In place of maxpooling, the convolutional layer with a Padding of 2, a Stride of 1, and a 3 3size filter is utilised, as demonstrated in the aforementioned example. During the step that involves downsampling, a convolution layer is used. This layer has twice as many feature channel filters as the layer that came before it, and the filter size and stroke for this layer are both set to 2. A fourth residual unit is utilised right at the very end of the contracting path in order to connect the two separate paths. In the same manner, three of the blocks that are still available are utilised in the construction of the expanding path. Following an operation that increases the size of the feature map by a factor of two, the next step is to perform a convolution using a filter of size 22 and create a chain of the corresponding feature maps of the contracting path. This procedure is executed before each block, and it runs in the order listed above. Convolution with a filter size of one and the Softmax activation function are used in the very last layer of the expanding path to divide the multi-channel feature maps into the optimal number of classes.

The vanishing gradient problem manifests itself during the training process for deep CNN networks [16]. As training continues, the gradient norms of earlier layers will eventually converge to zero. This can be seen as a measure of how well the layers are learning. The solution to this issue is a form of the learning strategy known as the residual network, which is referred to as ResNet. The output of one residual layer is combined with the input of the layer that comes before it in the ResNet algorithm. The result of this is that the input of one layer becomes the output of the layer that follows it.

Figure 4 illustrates how a residual learning block can be constructed by making use of Hx, which represents the residual mapping. This ResNet building block is getting us closer and closer to the correct answer, which is Hx 14 Fx x. Feed-forward neural systems that make use of "shortcut connections" are able to comprehend what Fx x is trying to say. In order to connect the input and output of a stacked layer without the requirement of any additional parameters, the shortcut connections make use of a method that is known as "identity mapping." Because of this, gradients can flow back easily, which results in a significant increase in the number of layers and a significant acceleration in the training process. A revised version of ResNet that incorporates identity connections in an effort to make the network even more useful than it already is currently under consideration. The gradient is back-propagated to earlier layers in a more straightforward manner using this method, which makes the training procedure more manageable. The Identity block and the Convolutional block are the two components of the ResNet model that are considered to be of the utmost significance.

In this study, each voxel is represented by a single isochromat, just as Rieger and his colleagues did in the study that they conducted in 2017. In the process of creating the dictionary, the discrete matrix form of the Bloch equations is utilised [17]. It is assumed that the effect of the RF pulses is instantaneous and that the RF slice profile provides the optimal results in order to keep things as straightforward as possible. In the model, the various levels of B1+ are represented by a scaling factor that is added to the nominal FAs. These FAs are the starting point for the analysis. The current implementation simulates the following ranges based on our target field strength of 3T: T1 = [100:20:4000] milliseconds, T2 = [5:5:30 32:2:130, 135:5:200, 210:10:350] milliseconds, and B1+ = [0.5:0.05:1.5]. The values that were selected for T1, T2, and T2 were chosen in order to focus on a particular range of brain tissue. The entire dictionary has somewhere in the neighbourhood of 350,000 entries in it [18]-[21]. In

addition, a more limited validation data set was produced in order to test the deep learning model while it is still in the process of being trained and to check for overfitting. The entries T1, T2, and B1 from the training set have been modified in this validation data set to ensure that there is no overlap between the two dictionaries. There are 33,000 entries total in the validation data set. T1 has the same value as [110:40:4000], T2 has the same value as [11:20:350], and B1 has the same value as [0.525:0.05:1.5]. On a desktop computer with a processor running at 1.6 gigahertz, it is possible to compile a dictionary in significantly less than five minutes.

2.1. Problem Statement

Tumours, in general, and brain tumours in particular, are well-known for their potential danger to our health. It is possible for a patient's brain tumour to be of any size, shape, or colour. As a result, locating and diagnosing tumours can be difficult. To determine whether or not an image depicting an abnormality in the brain contains an MRI tumour, a person's level of expertise and other factors must be taken into consideration. The use of MRI images to detect brain tumours can help solve problems and produce better outcomes. Medical professionals and those in charge of treatment planning face a challenge in detecting brain tumours in patients based on their various symptoms.

3. Conclusion

This research is a comprehensive meta-review that includes reviews as well as surveys pertaining to deep learning in the medical field. A comprehensive search was conducted with the aid of the widely utilised medical database PubMed, and the results revealed the existence of more than forty recent reviews and surveys. In the process of developing a treatment strategy, the segmentation of any kind of tumour is a very important step. Deep Neural Networks are useful tools that can be utilised during the segmentation process. However, they have to contend with gradients that become progressively less severe as the training progresses. In this piece of writing, the Residual Network, also known as ResNet, is discussed as a potential solution to the issue at hand. Because ResNet implements something called a "identity shortcut connection," the gradient can be backpropagated to earlier layers. It takes a lot less time to compute than other CNN, FCN (U-Net), and Unet-res methods, and the proposed method is more accurate than all of those methods combined.

References

[1] Shen, Dinggang, Guorong Wu, and Heung-Il Suk. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19: 221.

[2] Havaei, Mohammad, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. (2017). Brain tumour segmentation with deep neural networks. Medical Image Analysis, 35: 18-31.

[3] Jin, Qiangguo, Zhaopeng Meng, Changming Sun, Hui Cui, and Ran Su. (2020). RA-UNet: A hybrid deep attention-aware network to extract liver and tumour in CT scans. Frontiers in Bio. and Biotechnology, 1471.

[4] Bhalodiya, Jayendra M., Sarah N. Lim Choi Keung, and Theodoros N. Arvanitis. (2022). Magnetic resonance image-based brain tumour segmentation methods: A systematic review. Digital Health, 8: 20552076221074122.

[5] Abadi, Martin, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. (2016). Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 308-318.

[6] Yanase, Juri, and Evangelos Triantaphyllou. (2019). A systematic survey of computer-aided diagnosis in medicine: Past and present developments. Expert Systems with Applications, 138: 112821.

[7] Moran, Paul R., R. A. Moran, and N. Karstaedt. (1985). Verification and evaluation of internal flow and motion. True magnetic resonance imaging by the phase gradient modulation method. Radiology 154(2): 433-441.

[8] Taylor, Russell H., Arianna Menciassi, Gabor Fichtinger, Paolo Fiorini, and Paolo Dario. (2016). Medical robotics and computer-integrated surgery In Springer handbook of robotics, Springer, Cham, pp. 1657-1684.

[9] Prasad, Amit. (2005). Making images/making bodies: Visibilizing and disciplining through magnetic resonance imaging (MRI). Science, Technology, & Human Values, 30(2): 291-316.

[10] Sundaresan, S., K. Suresh, V. Kishore, and A. Jayakumar. (2021). Insight Into Various Algorithms For Medical Image Analyzes Using Convolutional Neural Networks (Deep Learning). In Handbook of Deep Learning in Biomedical Engineering and Health Informatics, Apple Academic Press, pp. 137-164.

[11] Liu, Li, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen. (2020). Deep learning for generic object detection: A survey. International Journal of Computer Vision, 128(2): 261-318.

[12] Tsai, Yi-Hsuan, Wei-Chih Hung, Samuel Schulter, Kihyuk Sohn, Ming-Hsuan Yang, and Manmohan Chandraker. (2018). Learning to adapt structured output space for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7472-7481.

[13] Li, Zhizhong, and Derek Hoiem. (2017). Learning without forgetting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(2): 2935-2947.

[14] Kumar, T. Ananth, G. Rajakumar, and TS Arun Samuel. (2021). Analysis of breast cancer using grey level co-occurrence matrix and random forest classifier. International Journal of Biomedical Engineering and

Technology, 37(2): 176-184.

[15] Mercioni, Marina Adriana, and Stefan Holban. (2020). The most used activation functions: Classic versus current. In 2020 IEEE International Conference on Development and Application Systems (DAS), pp. 141-145.

[16] Roodschild, Matías, Jorge Gotay Sardiñas, and Adrián Will. (2020). A new approach for the vanishing gradient problem on sigmoid activation. Progress in Artificial Intelligence, 9(4): 351-360.

[17] McGivney, Debra F., Eric Pierre, Dan Ma, Yun Jiang, Haris Saybasili, Vikas Gulani, and Mark A. Griswold. (2014). SVD compression for magnetic resonance fingerprinting in the time domain. IEEE Transactions on Medical Imaging, 33(12): 2311-2322.

[18] Pavithra, M., R. Rajmohan, T. Ananth Kumar, and S. G. Sandhya. (2021). An Overview of Convolutional Neural Network Architecture and Its Variants in Medical Diagnostics of Cancer and Covid-19. Handbook of Deep Learning in Biomedical Engineering and Health Informatics, pp. 25-49.

[19] Suresh, Kumar K., S. Sundaresan, R. Nishanth, and Kumar T. Ananth. (2021). Optimization and Deep Learning–Based Content Retrieval, Indexing, and Metric Learning Approach for Medical Images. Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications, pp. 79-106.

[20] Khajehim, Mahdi, Thomas Christen, Fred Tam, and Simon J. Graham. (2021). Streamlined magnetic resonance fingerprinting: Fast whole-brain coverage with deep-learning based parameter estimation. Neuroimage, 238: 118237.

[21] Acharya, U. Rajendra, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Muhammad Adam. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. Information Sciences, 415: 190-198.